University of Groningen

Comorbidities and medical history essential for mortality prediction in critically ill patients

Forte, Jose Castela; van der Horst, Iwan C. C.

Published in:
Lancet digital health

DOI:
10.1016/S2589-7500(19)30030-5

Comorbidities and medical history essential for mortality prediction in critically ill patients

See Articles page e78

Critically ill patients are a highly heterogenous population who tend to have many comorbidities. Often, patients admitted to intensive-care units (ICUs) with the same diagnosis and similar risk profiles according to available risk prediction scores have completely different clinical trajectories and outcomes. Even with increasingly large amounts of electronic health record data available, including clinical notes, vital sign measurements, laboratory data, and imaging data, the goal of unravelling complex disease mechanisms to better forecast patient outcomes remains largely unattained in critical care.[1]

Motivated by this problem, in *The Lancet Digital Health*, Annelaura Nielsen and colleagues[2] present the results of an innovative, exploratory analysis predicting in-hospital, 30-day, and 90-day mortality on the basis of a large and uniquely detailed cohort of patients in ICUs. In addition to laboratory data and other clinical parameters obtained during the first 24 h of an ICU stay for more than 10 000 patients, this dataset also included detailed, 10-year medical histories before ICU admission for more than 230 000 individuals. Factors present before ICU admission, such as comorbidities and medical history, have long been known to affect the risk of future complications or chance of survival.[3] However, even previous machine learning efforts that included broad health record data paid insufficient attention to these factors,[1,4] and Nielsen and colleagues' study is the first to link detailed medical history data from a highly heterogeneous patient population to clinical parameters measured during ICU stays. Remarkably, the authors concluded that a simple feed-forward neural network model including only age, sex, and patients' previous 10-year disease history performed similarly (in terms of prediction of mortality risk) to the two most commonly used ICU risk scores (the Simplified Acute Physiology Score II and the Acute Physiologic Assessment and Chronic Health Evaluation II), and that the combination of medical history and comorbidities with high-frequency ICU data outperformed both scores (Matthews correlation coefficient 0·391 for in-hospital mortality *vs* 0·347 with the Simplified Acute Physiology Score II and 0·300 with the Acute Physiologic Assessment and Chronic Health Evaluation II).

Medical history and comorbidity data are very important for predictions of survival in patients in critical care—and especially for efforts to increase the applicability of these models in clinical and research settings. Risk prediction can inform decisions, but an ideal decision support system would need to be dynamic and informative. The Artificial Intelligence Clinician, an algorithm that generates actual treatment decisions or suggestions, is an example of what decision support systems in the ICU could be.[5] When the algorithm successfully decreased sepsis-related mortality in an independent cohort in silico,[5] debate was sparked about the steps that should be taken to enable similar reinforcement learning models to be applied clinically.[6,7] Reinforcement learning models are developed on the basis of historical data for previous decisions made by clinicians.[5–7] Therefore, to generate good treatment decisions, all data used in clinicians' decision-making processes should be included to prevent confounding.[7] Additionally, after beneficial decisions are generated, any clinical application should be preceded by a clear mapping of the causal links that help clinicians to interpret the reasoning behind the decision. Both complete data collection and the identification of these causal links are notoriously difficult when observational data are used, because these data are often initially collected for a different purpose (ie, research or clinical). Therefore, some data used in clinicians' decision making might be missing, either because they were hard to identify or even unmeasurable, or because they were simply not included in the analysis despite being obtainable, making the dataset inappropriate for adjustment.[7–9]

Extensive patient characterisation is essential to maximise data quality, and, subsequently, the methodological correctness and clinical utility of machine learning models. However, even if all possibly relevant data were identified to minimise confounding, definition of outcomes of interest and strategies to gather these data prospectively can be equally challenging. As Gottesman and colleagues emphasised, focusing on short-term outcomes remains a challenge even for topics that are already broadly studied in critical care.[7] Studies focused on long-term outcomes, however, face a different type of problem. The difficulty with defining short-term targets for critical illnesses stems from the intricacy of the pathophysiology of these illnesses, which is undoubtedly a major issue in critical care, but one that will probably be solved with further research.[10] Long-term outcomes are different in that they relate to the less obvious core goal of critical care: healthy recovery after an acute ICU admission. Gathering the data necessary for research focused on long-term outcomes will require changes to data collection strategies and infrastructure, including closer collaboration between clinicians, researchers, and data scientists, and national medical data registries (appendix).

See Online for appendix

Overall, Nielsen and colleagues provide captivating evidence for the inclusion of comorbidities and medical history in mortality prediction models for ICU patients.[3] However, their findings also contribute to a broader debate that extends beyond predictive modelling, which was prompted by advances in machine learning research in critical care, and increasing awareness of heterogeneity in treatment response and issues with long-term patient-centered outcomes. It is clear that interpretability and trustworthiness need to be achieved before decision support systems for prediction and decision policy recommendations can be applied in clinical contexts. The causal links between predictions, outcomes, and automated policy recommendations will have to be studied further, starting with the exploration of detailed comorbidity and medical history data.

*José Castela Forte, *Iwan C C van der Horst*
Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Groningen, Netherlands (JCF); and Department of Critical Care, University Medical Center Groningen, Groningen, Netherlands (JCF, ICCvdH)
i.c.c.van.der.horst@umcg.nl

We declare no competing interests.

1 Marafino BJ, Park M, Davies JM, et al. Validation of prediction models for critical care outcomes using natural language processing of electronic health record data. *JAMA Netw Open* 2018; **1:** e185097.
2 Nielsen AB, Thorsen-Meyer H-C, Belling K, et al. Survival prediction in intensive-care units based on aggregation of long-term disease history and acute physiology: a retrospective study of the Danish National Patient Registry and electronic patient records. *Lancet Digital Health* 2019; **1:** e78–89.
3 Beck MK, Jensen AB, Nielsen AB, Perner A, Moseley PL, Brunak S. Diagnosis trajectories of prior multi-morbidity predict sepsis mortality. *Sci Rep* 2016; **6:** 36624.
4 Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med* 2018; **1:** 1609.
5 Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med* 2018; **24:** 1716–20.
6 Jeter R, Josef C, Shashikumar S, Nemati S. Does the "Artificial Intelligence Clinician" learn optimal treatment strategies for sepsis in intensive care? *arXiv* 2019; published online Feb 8. DOI:1902.03271 (preprint).
7 Gottesman O, Johannson F, Meier J, et al. Evaluating reinforcement learning algorithms in observational health settings. *arXiv* 2018; published online May 31. DOI:1805.12298 (preprint).
8 Bareinboim E, Pearl J. Causal inference and the data-fusion problem. *PNAS* 2016; **113:** 7345–52.
9 Pearl J. The seven tools of causal inference, with reflections on machine learning. *Commun ACM* 2019; **62:** 54–60.
10 Hernández G, Ospina-Tascón GA, Damiani LP, et al. Effect of a resuscitation strategy targeting peripheral perfusion status *vs* serum lactate levels on 28-day mortality among patients with septic shock: the ANDROMEDA-SHOCK randomized clinical trial. *JAMA* 2019; **321:** 654–64.